(54) Title: COGNITIVE FLOW CONTROL BASED ON CHANNEL QUALITY CONDITIONS
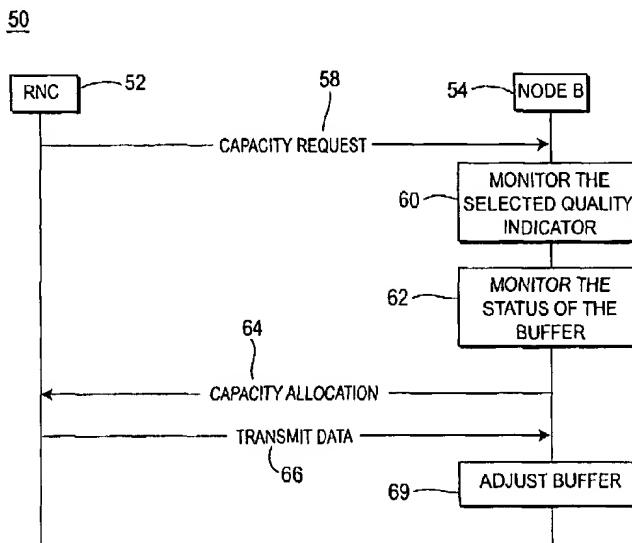
(57) Abstract: A system and method which improve the performance of a wireless transmission system by intelligent use of the control of the flow of data between a radio network controller (RNC) and a Node B. The system monitors certain criteria and, if necessary, adaptively increases or decreases the data flow between the RNC and the Node B. This improves the performance of the transmission system by allowing retransmitted data, signaling procedures and other data to be successfully received at a faster rate, by minimizing the amount of data buffered in the Node B. Flow control is exerted to reduce buffering in the Node B upon degradation of channel qualities, and prior to a High Speed Downlink Shared Channel (HS-DSCH) handover.

[0001]   COGNITIVE FLOW CONTROL
               BASED ON CHANNEL QUALITY CONDITIONS

[0002]   FIELD OF INVENTION

[0003]   The present invention relates to the field of wireless communications. More specifically, the present invention relates to the exertion of flow control for data transmissions between a radio network controller (RNC) and a Node B in a third generation (3G) telecommunication system.

[0004]   BACKGROUND

[0005]   A 3G Universal Terrestrial Radio Access Network (UTRAN) comprises several RNCs, each of which is associated with one or more Node Bs, and each Node B being further associated with one or more cells.

[0006]   The 3G Frequency Division Duplex (FDD) and Time Division Duplex (TDD) modes typically use the RNC to distribute (i.e., buffer and schedule), data transmissions to at least one User Equipment (UE). However, for the high speed channels of a 3G cellular system, data is scheduled for transmission by the Node B. One of these high speed channels, for example, is the High Speed Downlink Shared Channel (HS-DSCH). Since data is scheduled by the Node B, it is necessary to buffer data in the Node B for transmission to the UE(s).

[0007]   There are many scenarios where large amounts of data buffered in the Node B have a negative impact on the overall operation of the system. Several of these scenarios will be described hereinafter.

[0008]   The first scenario is related to the retransmission mechanisms in 3G systems to achieve high reliability of end-to-end data transmissions. It would be understood by those of skill in the art that the transmission failure between the Node B and the UE could be due to many different reasons. For example, the Node B may have retried the transmission several times without success. Alternatively, the transmission time allotted for a particular transmission may have expired. The present invention which will be described in further detail

-1-

hereinafter is intended to cover both these situations and any other situations where the failure of a data transmission necessitates a radio link control (RLC) retransmission.

[0009]   There are many levels of retransmission mechanisms.   One mechanism, for example, is the retransmissions of the Hybrid Automatic Repeat Request (H-ARQ) process for High Speed Downlink Packet Access (HSDPA). The H-ARQ process provides a mechanism where transmissions that are received in error are indicated to the transmitter, and the transmitter retransmits the data until the data is received correctly.

[0010]   In addition to the H-ARQ process, there are entities in the RNC and the UE. The sending RLC entity signals a sequence number (SN) in the header of a particular protocol data unit (PDU) which is used by the receiving RLC entity to ensure that no PDUs are missed in the transmission. If there are PDUs missed during the transmission, as realized by an out-of-sequence delivery of PDUs, the receiving RLC entity sends a status report PDU to inform the sending RLC entity that certain PDUs are missing. The status report PDU describes the status of successful and/or unsuccessful data transmissions. It identifies the SNs of the PDUs that are missed or received. If a PDU is missed, the sending RLC entity will retransmit a duplicate of the missed PDU to the receiving RLC entity.

[0011]   The impact of retransmissions in system performance will be described with reference to Figure 1. As shown, when the PDU with SN = 3 is not received successfully by the UE, the RLC within the UE requests its peer entity in the RNC for a retransmission. In the interim, the PDUs with SNs = 6 and 7 are queued in the buffer of the Node B.

[0012]   Referring to Figure 2, since the retransmission process takes a finite amount of time and data is continuing to be transmitted, two more PDUs with SNs = 8 and 9 have queued up behind the PDUs with SNs = 6 and 7, and in front of the retransmitted PDU with SN = 3. The PDU with SN = 3 will have to wait until the PDUs with SNs = 6-9 have been transmitted to the UE. Additionally, due to the requirement of in-sequence delivery of data to higher layers, the PDUs

-2-

with SNs = 4-9 will not be passed through higher layers until the PDU with SN = 3 is received and in-sequence delivery of data can be performed.

[0013] The UE will be required to buffer the out-of-sequence data until the missing PDU can be transmitted. This not only results in a delay of the transmission, but requires the UE to have a memory capable of data buffering for continuous data reception until the missed data can be successfully retransmitted. Otherwise, the effective data transmission rate is reduced, thereby affecting quality of service. Since memory is very expensive, this is an undesirable design constraint. Accordingly, this first scenario is when there is a need for RLC retransmission and a large amount of data buffered in the Node B results in a larger data retransmission delay and higher UE memory requirements.

[0014] A second scenario when the buffering of data in the Node B negatively affects system performance is in the case that layer 2 (L2) or layer 3 (L3) messages and data transmissions are processed by the same scheduling processes or share a single buffer in the Node B. While data is being buffered and processed and an L2/L3 message comes behind it, the message cannot circumvent the transmission queue. The greater the amount of data within a transmission buffer, (which operates as first-in-first-out (FIFO) buffer), the longer it takes for an L2/L3 message or data to get through the buffer. Any higher priority L2/L3 messages are thus delayed by the data in the buffers.

[0015] A third scenario where the buffering of data in the Node B could negatively impact the performance of the system is in the event of a serving HS-DSCH cell change. Since the Node B performs scheduling and buffering of data for an HS-DSCH, when the UE performs a serving HS-DSCH cell change from a source Node B to a target Node B, there is a possibility that considerable amounts of data may still be buffered in the source Node B after the handover. This data is not recoverable because there is no mechanism that exists within the UTRAN architecture for data buffered as the source Node B to be transmitted to the target Node B. Upon a serving HS-DSCH cell change, the RNC has no information regarding how much, if any, data is lost since the RNC it does not

-3-

know what data is buffered in the source Node B. The greater the amount of data that is buffered in the Node B in the event of an HS-DSCH cell change, the greater the amount of data which will ultimately be stranded in the source Node B and will have to be retransmitted.

[0016]   Accordingly, it would be desirable for the aforementioned reasons to limit the amount of data that is buffered in the Node B.

[0017]   SUMMARY

[0018]   The present invention is a system and method which improve the performance of a wireless transmission system by intelligent use of the control of the flow of data between the RNC and the Node B. The system monitors certain criteria and, if necessary, adaptively increases or decreases the data flow between the RNC and the Node B. This improves the performance of the transmission system by allowing retransmitted data, signaling procedures and other data to be successfully received at a faster rate than in prior art systems, by minimizing the amount of data buffered in the Node B. Flow control is exerted to reduce buffering in the Node B upon degradation of the channel quality, and prior to an HS-DSCH handover.

[0019]   In a preferred embodiment, the present invention is implemented in a wireless communication system including a radio network controller (RNC) in communication with a Node B having at least one buffer therein for storing data. The RNC signals the Node B with a request that the RNC send a certain amount of data to the Node B. The Node B monitors a selected quality indicator and calculates a capacity allocation for the buffer based on the selected quality indicator. The Node B signals the capacity allocation to the RNC. In response to receipt of the capacity allocation, the RNC transmits data to the Node B at a data flow rate determined in accordance with the capacity allocation.

[0020]   BRIEF DESCRIPTION OF THE DRAWING(S)

[0021]   A more detailed understanding of the invention may be had from

-4-

the following description, given by way of example and to be understood in conjunction with the accompanying drawings wherein:

[0022]   Figure 1 shows prior art buffering of data in the RNC, the Node B and the UE.

[0023]   Figure 2 shows prior art buffering of data in the RNC, the Node B and the UE in the event of a retransmission.

[0024]   Figures 3A and 3B, taken together, are a method in accordance with the present invention for monitoring the channel quality and adjusting the flow of data between the RNC and the Node B.

[0025]   Figure 4 is the buffering of data in the RNC, the Node B and the UE in the event of a retransmission using the method of Figures 3A and 3B.

[0026]   DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT(S)

[0027]   The present invention will be described with reference to the drawing figures wherein like numeral represent like elements throughout. Although the present invention will be described by referring to a specific number of PDUs being queued in a buffer (such as ten PDUs), this number of PDUs is referred to only for simplicity. The actual number of PDUs which is being transmitted and buffered in accordance with the aforementioned scenarios is more likely on the order of several hundred PDUs or more. The present invention and the teachings herein are intended to be applicable to any number of PDUs and any size transmission buffer.

[0028]   In general, the present invention reduces the flow of data to the Node B for a UE when there is a degradation of channel quality of the UE, and increases the flow of data to the Node B when there is an improvement in channel quality of the UE. In order to control the flow of the transmission of data between the RNC and the Node B, the present invention monitors one or more parameters for channel quality. This flow control can be based on one criterion, or a combination of many different criteria. Additionally, as will be explained in detail hereinafter, the criterion may be internally-generated by the Node B, or

-5-

may be generated by an external entity, (such as the UE), and sent to the Node B.

[0029]   Referring to Figure 3A, a method 50 in accordance with the present invention for monitoring the quality of a communication channel and adjusting the flow of data between the RNC 52 and the Node B 54 is shown. This method 50 handles the transmission of data between the RNC 52 and the Node B 54. The RNC 52 transmits a capacity request to the Node B 54 (step 58). The capacity request is basically a request from the RNC 52 to the Node B 54 that the RNC 52 would like to send a certain amount of data to the Node B 54. The Node B 54 receives the capacity request and monitors the selected quality indicator (step 60). This selected quality indicator may be based upon data transmitted from the UE (as will be described in detail hereinafter), or may be based upon an internally-generated quality indicator, such as the depth of the buffer in the Node B 54.

[0030]   The Node B 54 also monitors the status of the buffer within the Node B (step 62). As would be understood by those with skill in the art, although the present invention is described with reference to a single buffer within the Node B 54 for simplicity, most likely the buffer comprises a plurality of buffers or a single buffer segmented into a plurality of sub-buffers, each buffer or sub-buffer being associated with one or more data flows. Regardless of whether there is one or more multiple buffers, an indicator which indicates the amount of data within the buffer is generated internally within the Node B. This permits the Node B 54 to monitor the amount of data within the buffer, and also to monitor the amount of additional data the buffer may accept.

[0031]   The Node B 54 calculates and transmits a capacity allocation (step 64) to the RNC 52. The capacity allocation is an authorization by the Node B 54 to permit the RNC 52 to transmit a certain amount of data. The RNC 52, upon receiving the capacity allocation, transmits the data in accordance with the allocation (step 66). That is, the RNC 52 sends data to the Node B 54, the amount of which may not exceed the capacity allocation. The Node B then adjusts its buffer accordingly to receive and store the data (step 69). The amount

-6-

of data stored in the buffer will change in accordance with the incoming data that is transmitted from the RNC 52 and the outgoing data that is transmitted to the UE 82 (shown in Figure 3b).

[0032]   It would be appreciated by those of skill in the art that the method 50 shown in Figure 3A is constantly repeated as data flows from the RNC 52 to the Node B 54, and as the flow rate is continually adjusted by the Node B 54. It should also be noted that method steps 58, 60, 62, 64, 66 and 69 are not necessarily performed in sequence, and any one step may be applied multiple times before a different step in method 50 is applied. Additionally, some of the steps, such as the capacity allocation step 64, may indicate a repetitive data allocation that allows for the transmission of data (step 66) to be periodically implemented.

[0033]   Referring to Figure 3B, a method 80 in accordance with the present invention for monitoring the quality of a communication channel between the Node B 54 and a UE 82 is shown. The Node B 54 transmits data to the UE 82 (step 84). The UE 82 receives the data and transmits a signal quality indicator (step 86) such as the channel quality index (CQI) to the Node B 54. This signal quality indicator may then be used as the selected quality indicator in step 60 of Figure 3A.

[0034]   It would be noted by those of skill in the art that steps 84 and 86 are not necessarily sequential in practice. For example, in the FDD mode, signal quality indicators are periodically sent from the UE 82 regardless of whether or not a data is transmitted. In such a case, the UE 82 transmits a signal quality indicator either periodically or in response to a specific event to the Node B 54. This signal quality indicator may then be used as selected quality indicator in step 60 of Figure 3A.

[0035]   As aforementioned, the selected quality indicator may be internally generated by the Node B, or externally generated by another entity such as the UE and sent to the Node B. In accordance with a first embodiment, the criterion is the channel quality feedback from the UE. In this embodiment, the CQI which is an indicator of the downlink channel quality is used.

-7-

[0036]   In a second embodiment, the criterion is the ACK or NACK that the
UE produces in accordance with the H-ARQ process. For example, the number of
ACKs and/or the number of NACKs over a certain time period can be used to
derive an indication of the quality of the channel.

[0037]   In a third embodiment, the criterion is the choice by the Node B of
the modulation and coding set (MCS) that is needed to successfully transmit
data. As would be understood by those of skill in the art, a very robust MCS is
used when channel conditions are poor. Alternatively, a less robust MCS may be
utilized when the channel conditions are good and a large amount of data may be
transmitted. The choice of the most robust MCS set may be utilized as an
indicator of poor channel quality conditions, whereas the use of the least robust
MCS may signify that channel quality conditions are favorable.

[0038]   In a fourth embodiment, the criterion is the depth of the queue
inside the Node B transmission buffer(s). For example, if the Node B 54 buffer is
currently storing a large amount of data, it is an indicator that channel quality
conditions may be poor since the data is "backing up" in the Node B buffer. A
buffer which is lightly loaded may be an indicator that channel quality conditions
are good and the data is not backing up.

[0039]   In a fifth embodiment, the criterion is the amount of data that is
"dropped" in the Node B. As understood by those of skill in the art, the dropped
data is data which the Node B has attempted to retransmit several times and has
given up after a predetermined number of retries. If a large number of
transmissions are dropped by the Node B, it is an indicator that channel quality
conditions are poor.

[0040]   In a sixth embodiment, the criterion is the amount of data that can
be transmitted by the Node B within a predetermined duration, such as one
hundred milliseconds. Depending upon the quality of a communication channel,
the number of PDUs that are buffered in the Node B may change. Although the
predetermined duration may be fixed, due to changing channel quality conditions
the amount of PDUs that may be transmitted within the predetermined duration
may change dramatically. For example, if channel quality conditions are good, a

-8-

hundred PDUs may be able to be transmitted within a hundred millisecond duration; whereas if channel quality conditions are very poor, only ten PDUs may be able to be transmitted within the hundred second duration.

[0041]   It should be understood by those of skill in the art that other criteria which may directly or indirectly indicate the condition of the channel may be utilized in accordance with the present invention. Additionally, a combination of two or more of the above-described criteria may be utilized or weighted accordingly, depending upon the specific needs of the system users.

[0042]   Referring to Figure 4, the benefits of adaptively controlling the flow of data between the RNC and the Node B can be seen. This example is the scenario where a retransmission is required due to a failed transmission and the flow of data between the RNC and the Node B is decreased. As a result of the data flow decrease, only one additional PDU with SN = 8 is queued in front of the retransmitted PDU with SN = 3. The exertion of flow control as shown in Figure 4 reduces the latency of the retransmission of the PDU with SN = 3 as compared to the prior art handling of retransmissions as shown in Figure 2 where the PDUs with SNs = 8 and are queued in front of the PDU with SN = 3. Therefore, the PDU with SN = 3 can be retransmitted to the UE earlier. The in-sequence delivery requirement results in faster processing and delivery of PDUs 4 through 8 to higher layers.

[0043]   While the present invention has been described in terms of the preferred embodiment, other variations which are within the scope of the invention as outlined in the claims below will be apparent to those skilled in the art.

*   *   *

CLAIMS

What is claimed is:

1.   In a wireless communication system including a radio network controller (RNC) in communication with a Node B having at least one buffer therein for storing data, a method for controlling the flow of data between the RNC and the Node B, the method comprising:

(a) the RNC signaling the Node B with a request that the RNC send a certain amount of data to the Node B;

(b) the Node B monitoring a selected quality indicator;

(c) the Node B calculating a capacity allocation for the buffer based on the selected quality indicator;

(d) the Node B signaling the capacity allocation to the RNC;

(e) in response to receipt of the capacity allocation, the RNC transmitting data to the Node B at a data flow rate determined in accordance with the capacity allocation and at least one predetermined criterion; and

(f) the Node B adjusting the buffer accordingly.

2.   The method of claim 1 wherein the selected quality indicator is based upon data transmitted from at least one User Equipment (UE) to the Node B.

3.   The method of claim 2 wherein the selected quality indicator is a channel quality index (CQI) which indicates the quality of a downlink channel established between the UE and the Node B.

4.   The method of claim 2 wherein the at least one predetermined criterion is based on an ACK or NACK that the UE produces in accordance with a Hybrid Automatic Repeat Request (H-ARQ).

-10-

5.   The method of claim 1 wherein the data flow rate increases in response to an improvement in the selected quality indicator.

6.   The method of claim 1 wherein the data flow rate decreases in response to a degradation in the selected quality indicator.

7.   The method of claim 1 wherein the selected quality indicator is internally generated by the Node B.

8.   The method of claim 1 wherein the at least one predetermined criterion is based on a modulation and coding set (MCS) selected by the Node B.

9.   The method of claim 1 wherein the at least one predetermined criterion is based on the depth of a queue inside the buffer of the Node B.

10.   The method of claim 1 wherein the at least one predetermined criterion is based on the amount of data that is dropped in the Node B.

11.   The method of claim 1 wherein the at least one predetermined criterion is based on the amount of data that can be transmitted by the Node B within a predetermined duration.

12.   A wireless communication system for controlling the flow of data, the system comprising:
        (a) a Node B having at least one buffer therein for storing data; and
        (b) a radio network controller (RNC) in communication with the Node B, wherein:
                (i) the RNC signals the Node B with a request that the RNC send a certain amount of data to the Node B;
                (ii) the Node B monitors a selected quality indicator;

-11-

(iii) the Node B calculates a capacity allocation for the buffer based on the selected quality indicator;

(iv) the Node B signals the capacity allocation to the RNC;

(v) in response to receipt of the capacity allocation, the RNC transmits data to the Node B at a data flow rate determined in accordance with the capacity allocation and at least one predetermined criterion; and

(vi) the Node B adjusts the buffer accordingly.

13.   The system of claim 12 further comprising:

(c) at least one User Equipment (UE) in communication with the Node B, wherein the selected quality indicator is based upon data transmitted from the UE to the Node B.

14.   The system of claim 13 wherein the selected quality indicator is a channel quality index (CQI) which indicates the quality of a downlink channel established between the UE and the Node B.

15.   The system of claim 13 wherein the at least one predetermined criterion is based on an ACK or NACK that the UE produces in accordance with a Hybrid Automatic Repeat Request (H-ARQ).

16.   The system of claim 12 wherein the data flow rate increases in response to an improvement in the selected quality indicator.

17.   The system of claim 12 wherein the data flow rate decreases in response to a degradation in the selected quality indicator.

18.   The system of claim 12 wherein the selected quality indicator is internally generated by the Node B.
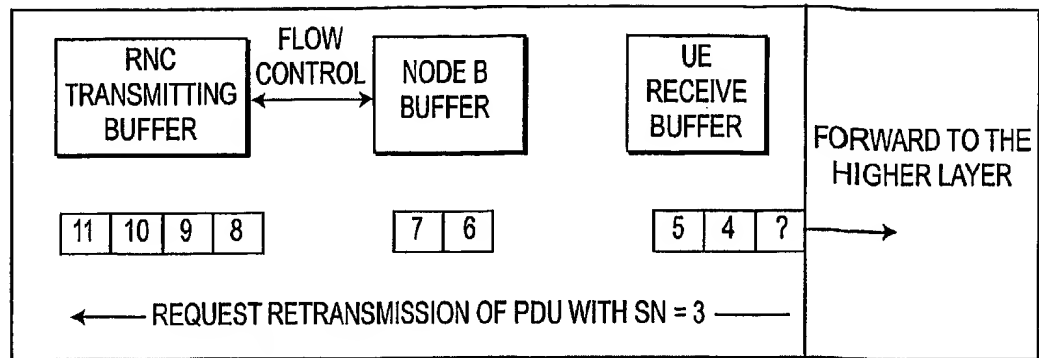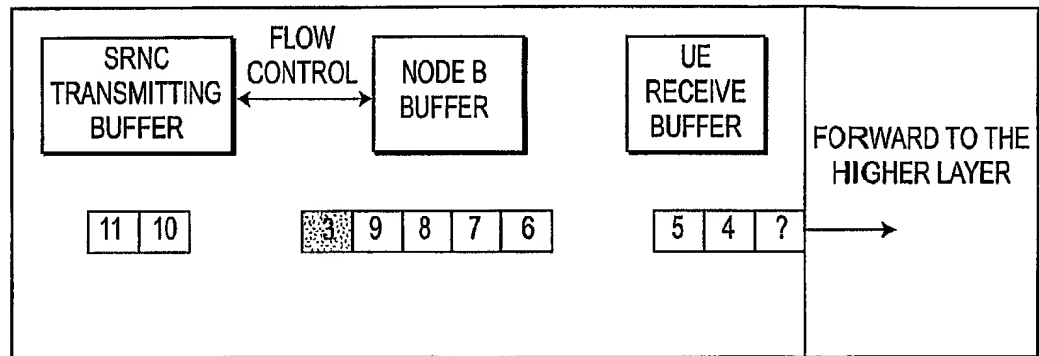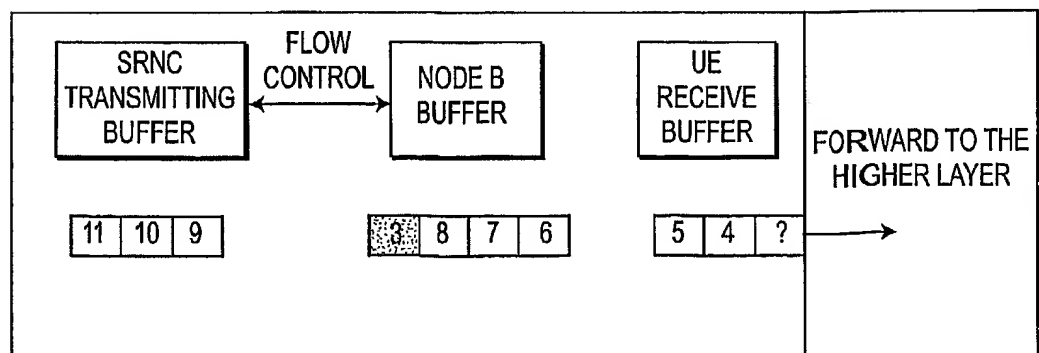

-12-

19.   The system of claim 12 wherein the at least one predetermined criterion is based on a modulation and coding set (MCS) selected by the Node B.

20.   The system of claim 12 wherein the at least one predetermined criterion is based on the depth of a queue inside the buffer of the Node B.

21.   The system of claim 12 wherein the at least one predetermined criterion is based on the amount of data that is dropped in the Node B.
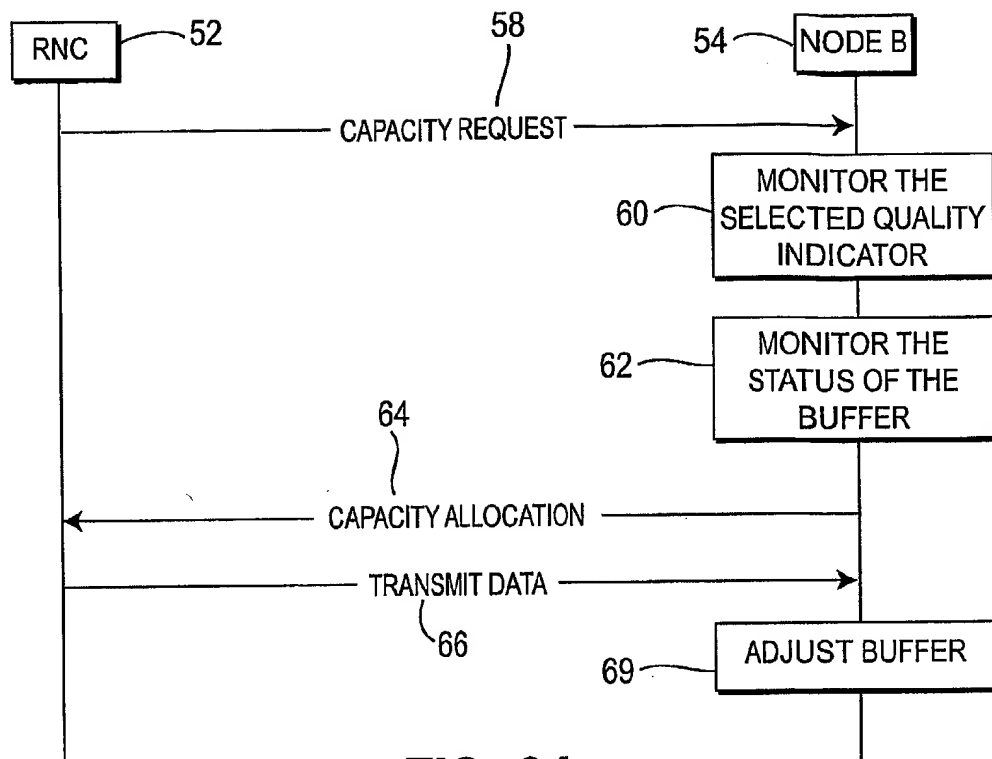
22.   The system of claim 12 wherein the at least one predetermined criterion is based on the amount of data that can be transmitted by the Node B within a predetermined duration.
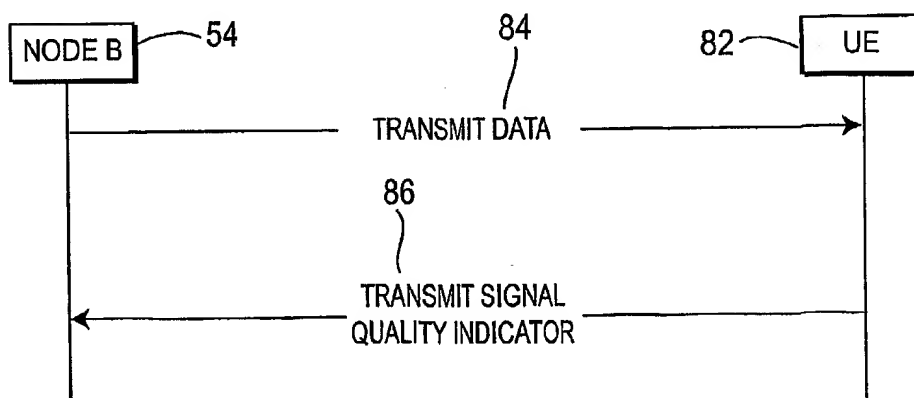
-13-

1/2

FIG. 1

FIG. 2

FIG. 4

50

RNC — 52   58   54 — NODE B

CAPACITY REQUEST

60 — MONITOR THE SELECTED QUALITY INDICATOR

62 — MONITOR THE STATUS OF THE BUFFER

64

CAPACITY ALLOCATION

TRANSMIT DATA

66

69 — ADJUST BUFFER

FIG. 3A

80

NODE B — 54   84   82 — UE

TRANSMIT DATA

86

TRANSMIT SIGNAL QUALITY INDICATOR

FIG. 3B